

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



100  
= 4

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> :	A1	(11) International Publication Number: WO 95/22818 (43) International Publication Date: 24 August 1995 (24.08.95)
(21) International Application Number: PCT/US94/14775 (22) International Filing Date: 22 December 1994 (22.12.94)		(81) Designated States: CA, CN, FI, JP, KR, RU, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).
(30) Priority Data: 08/198,750 17 February 1994 (17.02.94) US		Published <i>With international search report.</i>
(71) Applicant: MOTOROLA INC. [US/US]; 1303 East Algonquin Road, Schaumburg, IL 60196 (US).		
(72) Inventors: STEWART, Kenneth, A.; 915 Countryside Drive #116, Palatine, IL 60067 (US). KOTZIN, Michael, D.; 321 Fox Hill Drive, Buffalo Grove, IL 60089 (US).		
(74) Agents: SIKTO, Anthony, G. et al.; Motorola Inc., Intellectual Property Dept., 1303 East Algonquin Road, Schaumburg, IL 60196 (US).		
(54) Title: METHOD AND APPARATUS FOR GROUP ENCODING SIGNALS		
(57) Abstract		
<p>A code division multiple access (CDMA) communication system reduces system self-interference and enhances system capacity by making rate selection decisions for individual speech encoders in concert with other speech encoders. The system utilizes perceptually weighted error metrics (401) as input into a rate controller (404) which determines and provides selected rates (402) back to the encoders (105). The system provides optimum voice quality and system capacity in that it allows specific encoders to decrease their rate, which improves capacity, as necessary while allowing other encoders to maintain their rates. This prevents needless degradation in voice quality at those times when system capacity needs to be temporarily increased.</p>		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Iceland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LJ	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

## METHOD AND APPARATUS FOR GROUP ENCODING SIGNALS

5

### Field of the Invention

The present invention relates to communication systems utilizing code division multiple access (CDMA) techniques and more specifically to variable rate speech encoding methods for 10 reduction of system self-interference and enhancement of system capacity in point-multipoint multiple-access links with colocated digital speech encoding in such communication systems.

15

### Background of the Invention

In recent years a variety of techniques have been used to provide multi-user mobile communications within a limited available radio-frequency spectrum. These methods have 20 included frequency division multiple access (FDMA), time division multiple access (TDMA), and code division multiple access (CDMA) or, more usually, hybrids of these methods. All of these methods have been employed within the past decade in the design of commercial cellular telecommunications systems: 25 witness the use of FDMA in the North American AMPS system, FD/TDMA in the European Groupe Speciale Mobile (GSM) standard, and - more recently - the adoption of a direct sequence FD/CDMA approach by the United States Telecommunications Industry Association as embodied in its IS-95 standard. In the IS- 30 95 standard, subscribers share one of several wideband radio channels in the cellular band. Several proposals for so-called personal communications systems (PCS) are also being designed on similar FD/CDMA principles.

Almost all recent cellular and PCS systems have used digital speech coding and forward channel error correction as the physical layer for voice communication. More interesting in this context, is the use of voice activity detection (VAD) to recognize

5 the presence or absence of speech on the part of the either calling party. In the absence of speech, the speech encoder may instruct the modulator or transmitter to which it is linked to reduce its output power to zero, or transmit occasional packets of information describing only the background noise at the either  
10 user's location. Reducing the radio transmitter's duty cycle in this fashion provides the twin benefits of a reduction in power consumption (which increases battery life in the case of the mobile unit) and a reduction in interference between users sharing the same RF spectrum. Depending on the circumstances of the  
15 conversation, a reduction in transmitted power of between 40% and 65% can be achieved. The amount of power reduction is ultimately limited by the extent to which the degraded voice quality which accompanies significant VAD techniques is considered acceptable.

20 The possibility of power reduction is particularly important for CDMA systems. In such systems, user capacity is inversely proportional to the amount of system self-interference. In the TIA IS-95 FD/CDMA standard, the approach is slightly broadened by the use of a variable rate speech encoder in place of simple on-off  
25 or discontinuous transmission methods. In the IS-95 standard, the encoded speech is separated into 20 ms intervals which the speech encoder may elect to encode at a effective bit rate of 8000 bps, 4000 bps, 2000 bps, or 800 bps. Both the base-station to mobile station (forward) and mobile station to base-station (reverse) IS-95  
30 links exploit variable rate encoding. In the case of the forward link, mean transmit power is reduced by scaling down the output power as the encoded rate decreases. Channel symbol repetition allows symbol combining at the mobile receiver and hence

maintenance of the energy per symbol to noise power spectral density ratio which determines link performance. It should be noted that mean transmit power - and hence system self-interference - is reduced by a factor of four during 800 bps

5 transmission. By averaging over the aggregate voice activity for typical two-way conversations, it has been estimated that when using the standard speech encoding and voice activity detection algorithm defined in TIA standard IS-96 the mean transmit power will drop to around 41% of its nominal value. This has a  
10 significant effect on system forward link capacity.

In current implementations, however, of the IS-95 air interface standard and its associated IS-96 speech encoder standard, each forward voice link is encoded in isolation. That is, speech encoders make individual determinations of the minimum  
15 encoded rate required to maintain acceptable voice quality without regard to the other voice channels sharing the same RF spectrum. This requires that the rate-determination algorithm in each speech encoder should always minimize its encoded rate, even when the encoded rates of the other speech encoders sharing the  
20 same spectrum does not require this. For example, if all speech encoders sharing the same channel at a base-station should simultaneously seek to transmit at a low rate, the reduction in the total output power at the base-station means that each speech  
25 encoder could relax to the next higher rate at no risk to system capacity. Since minimizing the mean transmitted rate for a variable rate speech encoder requires that the voice quality be compromised, isolated speech encoding gives up voice quality needlessly. Also, as the CDMA system approaches capacity, and constraints are placed on the transmitted rate of each speech  
30 encoder in order to temporarily boost capacity, such constraints must be applied blindly, with all voice links subject to the same constraint irrespective of the effect on voice quality. This is wasteful, since it is known that voice quality depends on many

different factors and so opportunities will exist for reducing the rate of specific encoders with the least overall effect on the voice quality experienced within the sector/cell.

Thus a need exists for a method and apparatus for global  
5 speech encoding on the forward link of an FD/CDMA system by  
making rate selection decisions for individual speech encoders in  
concert with all other speech encoders feeding the same sector/cell  
and RF channel.

10

#### Brief Description of the Drawings

FIG. 1 generally depicts, in block diagram form, a prior art CDMA base-station transmitter.

15 FIG. 2 generally depicts, in block diagram form, the prior art rate determination apparatus specified in speech encoding standard TIA IS-96.

20 FIG. 3 generally depicts, in block diagram form, a codebook excited linear predictive speech encoder of the type utilized in the preferred embodiment and described in detail in speech encoding standard TIA IS-96.

FIG. 4 generally depicts, in block diagram form, the use of a supervising processor or rate controller to group speech encode in accordance with the invention.

25 FIG. 5 generally depicts a rate/quality table utilized by the rate controller as part of a number of algorithms for optimizing the overall speech quality of a sector/cell subject to a constraint on transmitted power from the sector/cell in accordance with the invention.

30 FIG. 6 generally depicts, in block diagram form, an apparatus for implementing the CDMA group encoding method in accordance with the invention.

FIG. 7 generally depicts, in block diagram form, a rate controller which may beneficially implement group encoding in accordance with the invention.

5

#### Detailed Description of a Preferred Embodiment

A code division multiple access (CDMA) communication system reduces system self-interference and enhances system capacity by making rate selection decisions for individual speech encoders in concert with other speech encoders. The system utilizes perpetually weighted error metrics (401) as input into a rate controller (404) which determines and provides selected rates (402) back to the encoders (105). The system provides optimum voice quality and system capacity in that it allows specific encoders to decrease their rate, which improves capacity, as necessary while allowing other encoders to maintain their rates. This prevents needless degradation in voice quality at those times when system capacity needs to be temporarily increased.

The preferred embodiment of the invention is described as it relates to a CDMA digital cellular communications system based on the Telecommunications Industry Association standards IS-95 and IS-96. It will be appreciated by one skilled in the art that the invention may be applied to any CDMA point-to-multipoint link (generally the forward link of a digital cellular system) in which self-interference reduction by variable rate speech encoding is to be applied. However, the technique discussed may be beneficially utilized in any communication system, and in fact is not restricted to communication systems. For example, the technique may be utilized where speech encoding occurs for storage in a memory means having limited memory space. In essence, the technique is applicable to any application where encoding (be it speech, video, data, etc.) is utilized and constraints related to the encoding (be it

power level, encoding quality, system capacity, memory space, etc.) are present.

The method and apparatus group encodes signals by accepting rate determination information from at least two encoders and determining the rate of at least one encoder based on the rate determination information for the at least two encoders. In the preferred embodiment, rate determination information is quality information relating reconstruction quality as a function of encoding rate (on a 20 ms segment-by-20 ms segment basis). The 10 quality information includes, but is not limited to, perceptual weighting error metrics generated by the analysis-by-synthesis speech encoders, signal-to-noise (S/N) ratio, segmented S/N, cepstral distance, an LPC distance measurement and a BARK spectral distance measurement, all of which are well known in the 15 art.

The determination of the rate of the at least one encoder is based on a threshold criterion (which is typically predetermined). In the preferred embodiment, the threshold criterion may include, but is not limited to, the total output power of the sector/cell to 20 which the encoders are assigned, the total output power of an adjacent sector/cell, the current power level of transmission by a serving base-station, the current data rate of the at least two encoders, the memory available in a memory means, the processing power available in a processing means and the 25 bandwidth available in a predetermined spectrum. Also in the preferred embodiment, the encoders are variable rate analysis-by-synthesis encoders. These encoders may encode signals including, but not limited to, speech signals, video signals and data signals.

FIG. 1 shows the high-level architecture of the forward link 30 of a CDMA base-station (102) designed for the preferred embodiment of the TIA IS-95 digital cellular radio standard. The base-station (102) of FIG. 1 performs, inter alia, variable rate speech encoding, forward error correction, forward link power

control, multiple access spreading, and modulation and transmission. In FIG. 1, several standard  $\mu$ -law encoded, multiplexed, 64 kbps pulse code modulated (PCM) T1 links (101) from the public switched telephone network (PSTN) (100) are  
5 brought to a demultiplexer (103). Each 64 kbps voice link (104) is then passed through a digital speech encoder (105). In a conventional implementation, the speech encoding function is performed by a number of general purpose digital signal processors (DSP's) such as the Motorola DSP56156 processor, ROM  
10 coded DSP's, or application specific integrated circuits (ASICs). Several such processors are generally grouped onto a single printed circuit board (although this is not necessary for the invention) which is then capable of processing a full T1 trunk of multiplexed voice channels. After speech encoding, error  
15 correction (106) is applied in the form of convolutional and cyclic codes, followed by BPSK baseband modulation (107), Walsh cover and short pseudo-noise (PN) sequence spreading (108), low-pass filtering (109), transmit power level adjustment (110) and power amplification (111), and finally transmission to the mobile station  
20 (113) (for simplicity, frequency shifting to RF is not shown).

A block diagram of the TIA IS-96 standard processing performed by the DSP or other device used to implement the speech encoder (105) is shown in FIG. 2. As shown, speech encoder (105) can be broken down into two main elements: rate determination and encoding. Consider first the rate determination function. In the IS-96 standard, each speech encoder (105) divides its associated PCM signal stream into contiguous 20 ms frames consisting of 160 samples of the source speech waveform. The power level of each frame (which is the  
25 zeroth lag  $R(0)$  of the autocorrelation function estimate of the frame produced by the autocorrelation estimator (200)) is fed to a bank of comparators (203) which establish which of three monotonic-increasing threshold levels the frame power exceeds.  
30

These levels are generated by 2nd order interpolation of a non-linear average of the power level of the speech signal formed by block (201). Note that all these processing steps are completely defined in TIA standard IS-96. If the current frame energy is less than the lowest of the three thresholds, the frame is declared an 1/8 rate frame; if the frame energy lies between the lowest and middle of the thresholds, the frame is declared a 1/4 rate frame; if it is between the middle and highest threshold the declared a 1/2 rate frame; and finally, if the frame energy exceeds the highest threshold level, the frame is declared full rate frame. This final step is performed by comparators (203) and decoder (204) to produce the selected rate (205).

The selected rate (205) is then input to the codebook excited linear predictive (CELP) speech encoding function (206) which forms a parametric description of the speech frame using the specified number of bits for that rate. In the preferred embodiment, the number of bits used to express the encoded parameters of an 1/8 rate frame is 16 (ignoring additional bits used for error correction/detection); for a 1/4 rate frame, 40 bits; for a 1/2 rate frame, 72 bits; and for a full rate frame, 160 bits. While CELP is depicted and discussed in the preferred embodiment, other encoding techniques such as, inter alia, waveform coding, linear predictive coding (LPC), sub-band coding (SBC), code excited linear prediction (CELP), stochastically excited linear prediction (SELP), vector sum excited linear prediction (VSELP), improved multi-band excitation (IMBE), and adaptive differential pulse code modulation (ADPCM) coding algorithms may likewise be beneficially employed.

For the purpose of clarity, it is necessary to describe in more detail the CELP speech encoding procedure. A high-level block diagram of the signal processing used in the CELP speech encoder of the preferred embodiment appears in FIG. 3. As shown in FIG. 3, an estimate (309) of the autocorrelation function of consecutive

20 ms frames of the 64 kbps speech signal (104) is first obtained (this is usually done in common with block (200) of the rate determination procedure). Next, solution of the so-called Normal Equations using, for example, the Schur Recursion (301), provides

5 the short term linear predictive (STP) filter coefficients (310). Often, the STP filter (303) is a lattice filter, and the STP coefficients are lattice filter reflection coefficients. After quantization (302) by line spectral pairing or some other robust quantization method, the STP coefficients are used to filter the speech signal. The  
10 resulting signal is next passed to the long term prediction (LTP) filter (313) and (in the case of a CELP linear predictive coder) the codebook search procedure. The LTP filter is generally a first order recursive filter whose feedback delay and gain are variable – they appear in FIG. 3 as LTP lag L (304) and LTP gain G (305). Encoding  
15 then proceeds by simultaneously adjusting the LTP lag and gain and the codebook index I (312) so that the square error at the output of the LTP filter is minimized. L, G, and I are then quantized (often using simple biased linear quantizers methods), and passed along with the STP coefficients to the error correction  
20 block. The performance of this analysis-by-synthesis procedure can be improved by weighting the error metric which is to be minimized by the human auditory frequency response. This is done with a perceptual weighting filter (307) which modifies the error metric (308) to emphasize those frequency components to  
25 which the human ear is most sensitive. One skilled in the art will appreciate that the perceptually weighted error metric is made available by almost all sophisticated analysis-by-synthesis speech encoders. The present invention, as mentioned above, is therefore not limited exclusively to CELP speech encoders.

30 With this background, group speech encoding in accordance with the invention may now be described. It is clear from FIG. 1 and FIG. 2 that, in the prior art, the encoded rate of each forward link speech encoder is determined in isolation. That is, the

encoded rate of each 64 kbps voice link is determined exclusively by signal processing that speech signal. Since the amount of self-interference (and hence the capacity) in the forward link of a CDMA system depends on the mean encoded rate of each encoder,

5 it is also clear that in order to operate at the maximum possible capacity, the rate determination algorithm of each speech encoder must be designed to always seek the minimum possible rate, since each encoder operates in isolation and has no knowledge of the total power (and hence system self-interference) being emitted at

10 the base-station antenna (112). Since speech quality must be sacrificed to achieve low mean encoded rates, this implies that overall system speech quality is unnecessarily sacrificed when the system is not at its maximum capacity – or equivalently, is not transmitting its maximum allotted power. Put another way,

15 isolated speech encoding allows the total instantaneous output power at the base-station to have a large variance.

Since, in many CDMA power control algorithms, a strict limit is placed on total emitted power from a cell or sector, the rate used to encode individual links must be kept unnecessarily low.

20 In addition, it is known that the perceptual quality of a digitally encoded voice link is dependent not only on the speech encoder being used, but also on factors such as the gender, accent, loudness of the speaker, and environmental factors such as type/levels of acoustic background noise. Thus, by encoding each link in

25 isolation, no recognition is made of situations where one link may be reduced in rate with a smaller loss in perceived overall voice quality than an equivalent reduction in rate on another link, and hence another speaker. Further, the current art embodied in TIA standard IS-96 makes no use of the perceptually-weighted

30 encoding error in performing rate determination.

The method shown in FIG. 4 can be used to overcome these deficiencies. In FIG. 4, each speech encoder (105) evaluates, for each 20 ms frame, the perceptually weighted error metric (401)

produced by encoding the speech frame at each of the four candidate rates (more than four rates may be possible in alternate embodiments). This information is then passed back to a supervising rate controller (404). Rate controller (404) then forms  
5 a rate/quality table similar to that of FIG. 5, which depicts the perceptually-weighted error produced by encoding at each of the candidate rates for each of the N speech encoders reporting to the rate controller.

A simple approach to optimizing the overall voice quality  
10 of the cell or sector starts by assuming that all N voice channels have equal transmit power. All of the encoders (105) are placed in the lowest candidate rate and the total transmit power P is calculated by rate controller (404). In this case, P is simply equal to the sum of the rate values for all N encoders, where the rate value  
15 for 1/8 rate is 1/8, for 1/4 rate is 1/4, and so on. Rate controller (404) then finds the largest entry in the rate/quality table corresponding to the current candidate rate for any of the N encoders. This is equivalent to identifying the encoder with the worst voice quality (i.e. the largest perceptually weighted error) for  
20 the current set of selected rates. The rate for that encoder is increased to the next highest rate, and P is recalculated. This process continues until P exceeds some total power threshold T at which time the procedure terminates. An improved approach would be to apply the procedure to rate/quality table entries which  
25 have been weighted by the transmit gain associated with each encoder. This would be extracted from power level block (110). It will be appreciated by one of ordinary skill in the art that the overall effect of this procedure is to reduce power by sacrificing the rate of those encoders which will suffer the least reduction in  
30 quality by operating at a lower rate.

A more complex approach would be as follows. Assume that, as above, the goal (i.e., the predetermined criterion) of the rate-reduction scheme during periods of high traffic loading is to

maintain the overall transmitted power to be less than some threshold T, where T is set according to the current load conditions. Define a global measure Q of speech quality for the sector/cell served by the base-station to be the sum of the 5 perceptual errors for the current set of selected rates for the N voice channels. Each encoder is initialized to encode at the maximum rate. Q is then evaluated and the corresponding transmitted power calculated using either the equal power assumption or the weighted transmit power method described 10 above.

A simplification of this method would occur where the rate controller (404) was not available, but where each DSP was encoding several voice links by time-sharing its available computational resources. In that case, the rate selection procedure 15 would be applied over the number of voice channels for which the DSP was performing encoding. FIG. 6 generally depicts an apparatus which may be used to implement this scenario. In FIG. 6, a single DSP (603), such as the Motorola DSP56156, communicates via a time-division multiplexed serial bus or a 20 conventional parallel address/data bus. Rate determination information and rate selections are passed between the controlling DSP (603) and the DSP's (602) used for speech encoding via bus (604). Alternatively, the controlling DSP (603) may be eliminated and one of the encoder DSP's (602) promoted to fulfill the global 25 rate controller function and speech encoding for one or more voice channels.

FIG. 7 generally depicts, in block diagram form, a rate controller (404) which may beneficially implement group encoding in accordance with the invention. Rate controller (404) comprises 30 means for accepting (700) rate determination information (401) from a plurality of encoders (105). In the preferred embodiment, rate determination information is quality information which includes a perceptually weighted error metric. Means for accepting

(700) has its output entering means for determining (703) which determines encoding requirements based on predetermined criterion. The predetermined criterion include those stated above as threshold criterion. The output of means for determining (703)

- 5 is input into means for adjusting (706) which adjusts the encoding rate for any encoder out of the plurality of encoders based on the rate determination information and the predetermined criterion. In a scenario where the predetermined criterion is total transmit power or available memory space, means for adjusting (706) will
- 10 typically increase the encoding rate for the encoder having the worst quality (based on a determination/calculation of either total transmit power or available memory space and a threshold) as described above. However, certain predetermined criterion, such as system capacity, may require means for adjusting (706) to
- 15 decrease the encoding rate for a particular encoder.

While the invention has been particularly shown and described with reference to a particular embodiment, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit

- 20 and scope of the invention.

What I claim is:

## Claims

1. A method of group encoding signals comprising the steps of:

5

accepting rate determination information from a plurality of encoders;

determining encoding requirements based on predetermined criterion; and

10 adjusting the encoding rate for any encoder out of the plurality of encoders based on the rate determination information and the predetermined criterion.

15 2. The method of claim 1 wherein the predetermined criterion further comprises either the total output power of the sector/cell to which the encoders are assigned, the total output power of an adjacent sector/cell, the current power level of transmission by a serving base-station, the current data rate of the at least two encoders, the memory available in a memory means, 20 the processing power available in a processing means and the bandwidth available in a predetermined spectrum.

3. An apparatus for group encoding signals, the apparatus comprising:

25

means for accepting rate determination information from at least two encoders; and

means, coupled to the means accepting, for determining the rate of at least one encoder based on the rate determination 30 information for the at least two encoders.

4. The apparatus of claim 3 wherein the at least two encoders further comprise analysis-by-synthesis encoders.

5. The apparatus of claim 4 wherein the rate determination information further comprises quality information generated by the at least two encoders.

5

6. The apparatus of claim 5 wherein the quality information further comprises perceptual weighting error metrics generated by the analysis-by-synthesis speech encoders, signal-to-noise (S/N) ratio, segmented S/N, cepstral distance, an LPC distance measurement and a BARK spectral distance measurement.

10

7. The apparatus of claim 3 wherein the means for determining the rate of the at least one encoder further comprises means for determining the rate of the at least first encoder based on a threshold criterion.

15

8. The apparatus of claim 7 wherein the threshold criterion further comprises either the total output power of the sector/cell to which the encoders are assigned, the total output power of an adjacent sector/cell, the current power level of transmission by a serving base-station, the current data rate of the at least two encoders, the memory available in a memory means, the processing power available in a processing means and the bandwidth available in a predetermined spectrum.

20

9. The apparatus of claim 3 wherein the at least two encoders further comprise at least two variable rate encoders.

25

10. The apparatus of claim 3 wherein the signals further comprise either speech signals, video signals or data signals.

30

1 / 5

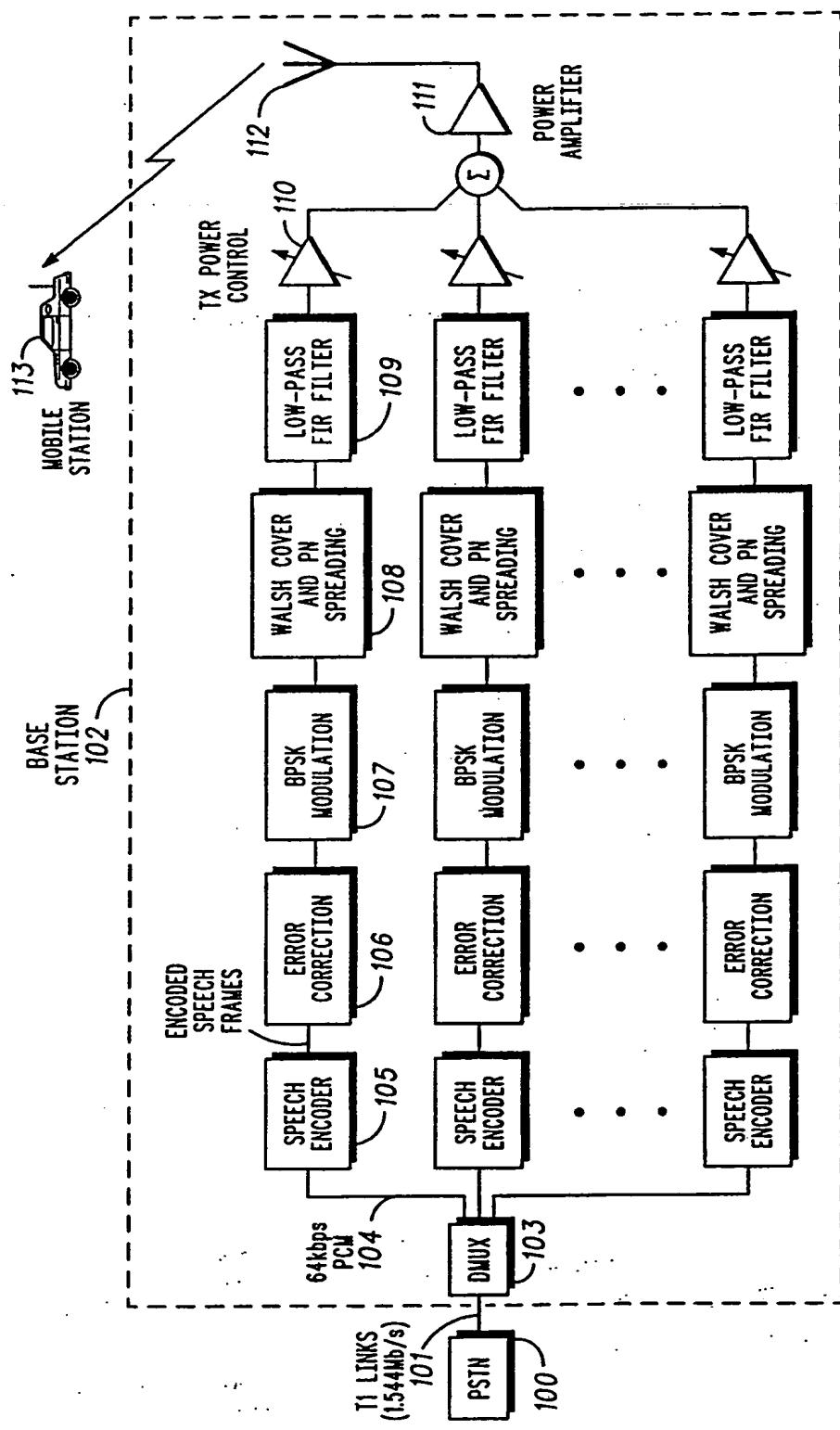


FIG. 1

2 / 5

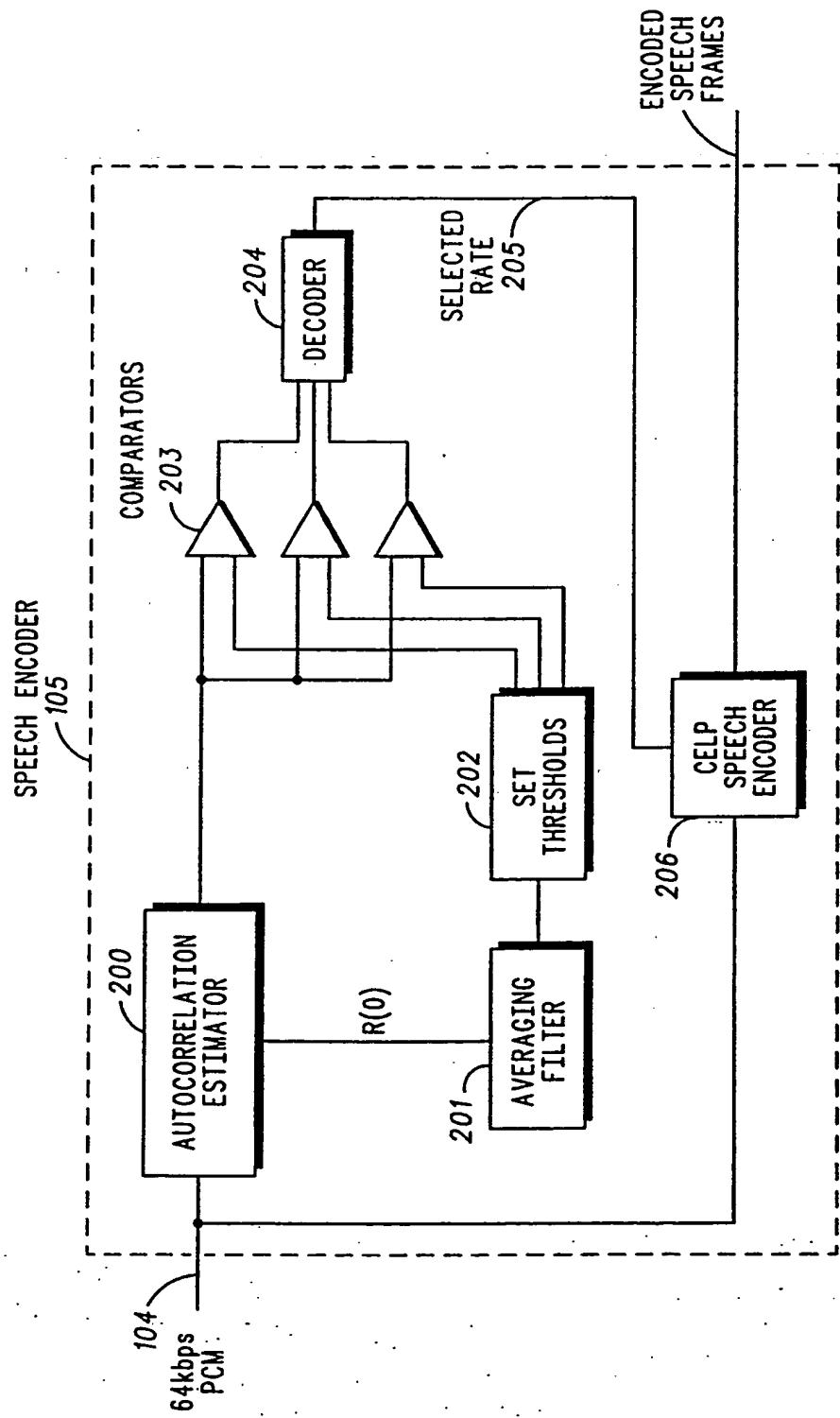
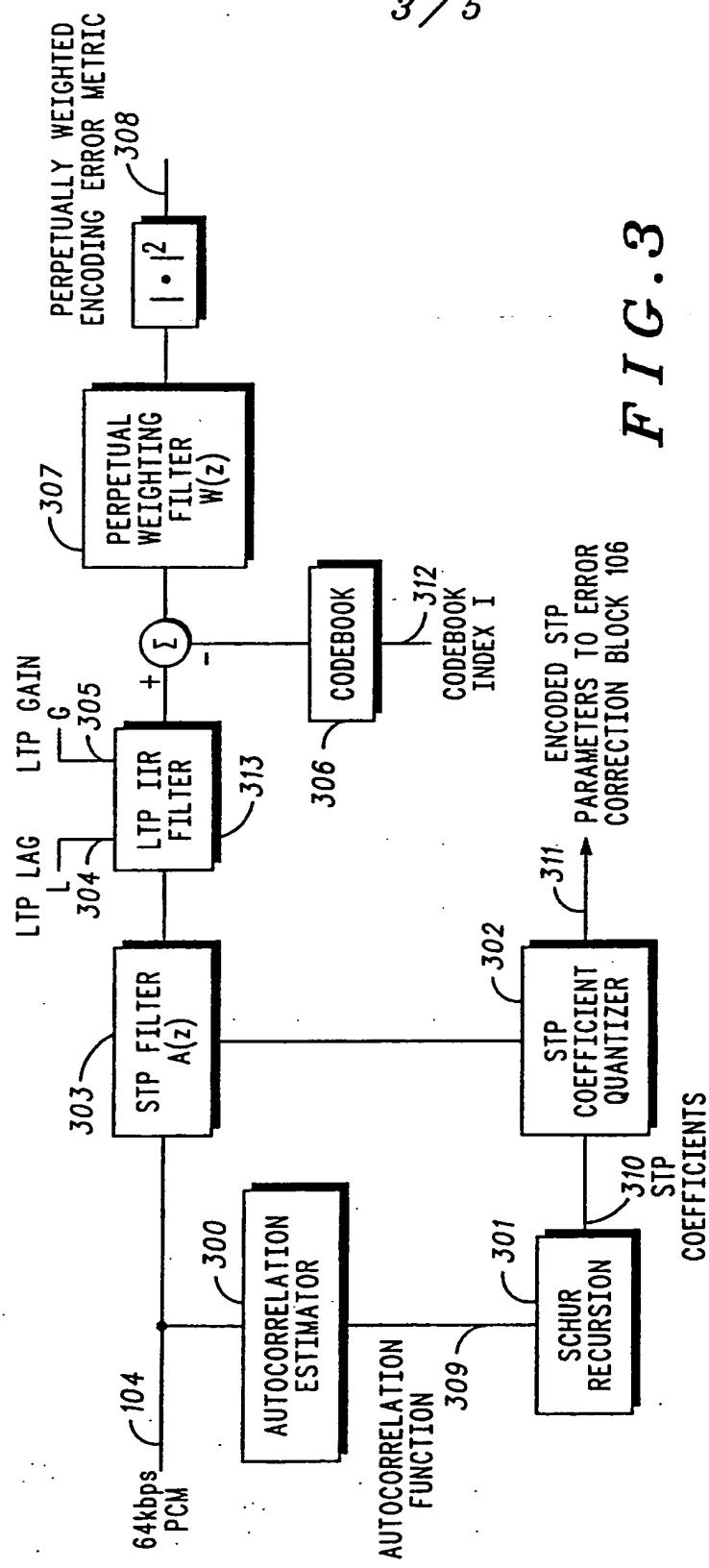


FIG. 2

3 / 5

**FIG. 3**

4 / 5

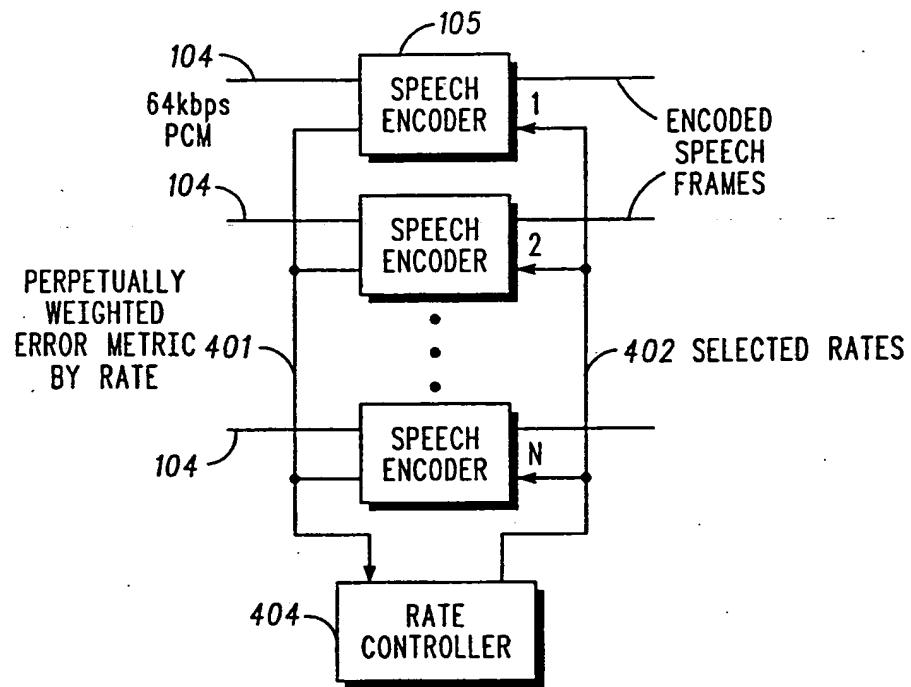


FIG. 4

CANDIDATE RATE				
	1/8	1/4	1/2	1/1
1	525	250	150	60
2	400	310	140	40
3	385	300	125	35
•	•	•	•	•
•	•	•	•	•
N	300	280	105	5

SPEECH ENCODER NUMBER

FIG. 5

5 / 5

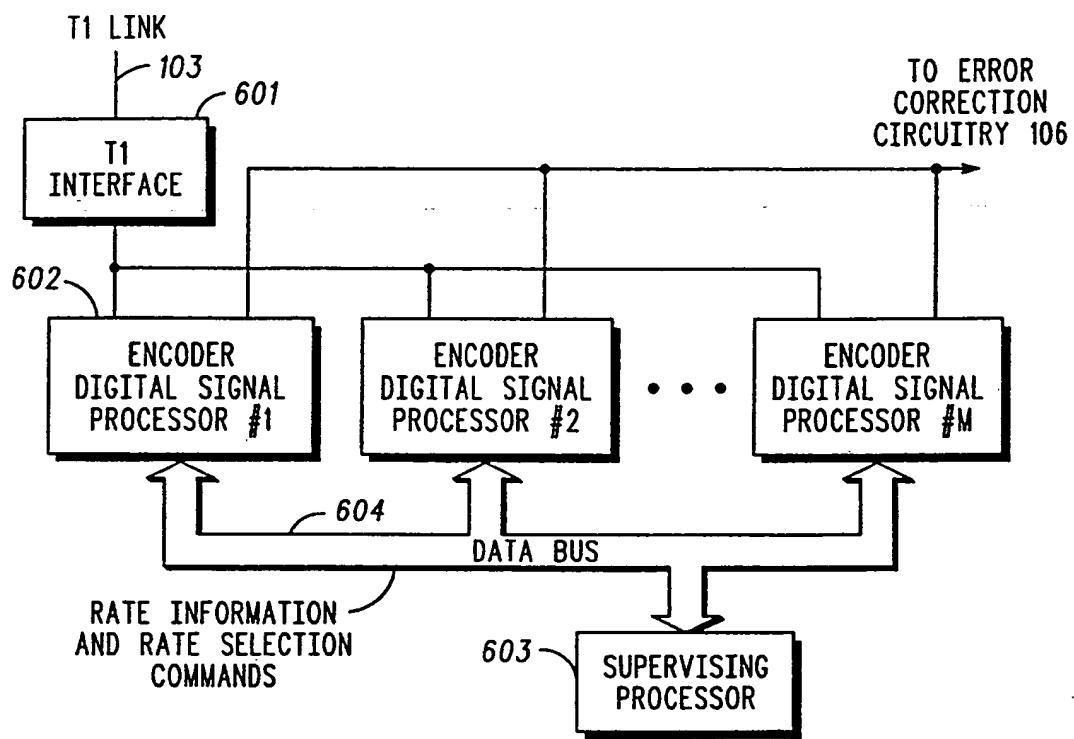


FIG. 6

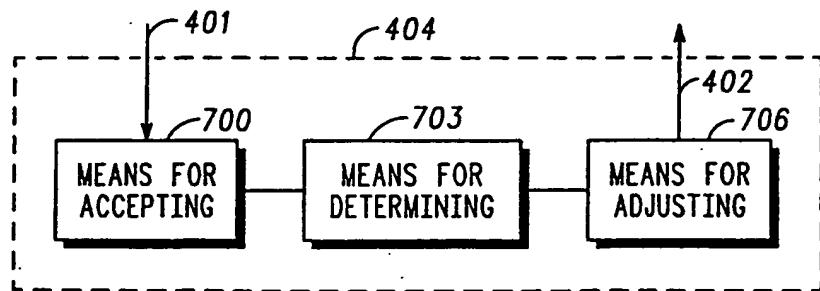


FIG. 7

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US94/14775

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :G10L 9/00  
US CL :395/2.29, 2.38, 2.39

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 395/2.28, 2.29, 2.3, 2.31, 2.32, 2.34, 2.38, 2.39; 381/36, 40

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US, A, 4,455,649 (ESTEBAN ET AL.) 19 June 1984, Fig. 2A, items 110F, 120F, 130F, 170, and 171; col. 2, line 64 - col. 3, line 10; col. 7, lines 8-15 and 34-38; col. 8, lines 1-4.	3, 5, 9-10 ----- 1-2, 4, 6-8
X	US, A, 5,150,387 (YOSHIKAWA ET AL.) 22 September 1992, Abs.; Fig. 1, items 104-1, 104-n, 108, and 109; Fig. 4, items 305 and 306; Fig. 6; col. 5, lines 34-58.	3, 5, 9-10 ----- 1-2, 4, 6-8

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means		
"P" document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

06 MARCH 1995

Date of mailing of the international search report

22 MAY 1995

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Faxsimile No. (703) 305-3230

Authorized officer  
*Michael Sartori*  
DR. MICHAEL A. SARTORI

Telephone No. (703) 305-9600

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US94/14775

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US, A, 4,949,383 (KOH ET AL) 14 August 1990, Fig. 1.	3, 5, 9-10 -----
Y		1-2, 4, 6-8